

LOGIQUES CLASSIQUE ET NON CLASSIQUES: LES FONDEMENTS THÉORIQUES DU RAISONNEMENT AUTOMATISÉ

BERNARD R. HODGSON

RÉSUMÉ. Les insuffisances de la logique classique (du premier ordre) dans la représentation de la connaissance et le raisonnement automatisé ont suscité l'intérêt pour l'emploi de logiques dites non classiques. Outre la présentation du cadre théorique général dans lequel la logique mathématique intervient en intelligence artificielle, le but du présent exposé est d'examiner de façon plus particulière comment le raisonnement automatisé trouve ses fondements dans les logiques tant classique que non classiques. Après un rappel de certains traits fondamentaux de la logique du premier ordre, nous présentons quelques approches récentes permettant d'effectuer des inférences dans des contextes imparfaits de connaissances, certaines de ces approches étant purement symboliques (logique modale, logiques non monotones) tandis que d'autres font appel à un formalisme numérique (logique probabiliste, logique possibiliste, logique floue).

ABSTRACT. The insufficiency of classical (first order) logic for knowledge representation and automated reasoning has raised interest for the use of so-called non classical logics. Besides the presentation of the general theoretical framework within which mathematical logic enters into play in artificial intelligence, the goal of this survey paper is to consider more specifically how foundations of automated reasoning are to be found both in classical as well as in non classical logics. After a brief review of the fundamentals of first order logic, we present some recent approaches which allow inferencing in certain contexts of imperfect knowledge, some of these approaches being purely symbolic (modal logic, nonmonotonic logics) while others are based on numeric formalisms (probabilistic logic, possibilistic logic, fuzzy logic).

1. Introduction. Sous l'influence en particulier des besoins grandissants de l'informatique, la logique mathématique connaît depuis plusieurs années un essor considérable, comme le soutient entre autres Nerode [Ne86], et ce tant dans ses développements théoriques que dans ses applications. Certains, tels Manna et Waldinger [MW8-], considèrent même la logique mathématique comme indispensable à l'informaticien professionnel, lui attribuant un rôle similaire en informatique à celui joué traditionnellement par le calcul différentiel et intégral en physique ou en génie.

C'est peut-être dans le domaine de l'Intelligence Artificielle (IA) que ce caractère fondamental de la logique ressort le plus clairement. Le désir de construire des machines manipulant des symboles de manière à reproduire certains aspects du raisonnement humain mène d'une part à la question de la *représentation de la connaissance* (comment exprimer des faits sous une forme appropriée à la machine) et d'autre part au problème de

Reçu le 10 juin 1990.

Exposé présenté dans le cadre de l'Ecole de mathématiques et d'intelligence artificielle qui s'est tenue à Nabeul (Tunisie) en mars 1988 et organisée par l'"International Council on Mathematics in Developing Countries" et l'Université de Tunis. L'auteur tient à remercier le Conseil de recherches en sciences naturelles et en génie du Canada pour son soutien financier. Durant la rédaction de cet article, l'auteur était professeur visiteur au Centre de recherches mathématiques de l'Université de Montréal.

©Association mathématique du Québec

l'élaboration de *structures d'inférence* (comment traiter ces représentations pour en tirer certaines conséquences): on retrouve là les deux ingrédients à la base même de la démarche de formalisation propre à la logique mathématique. Et s'il est vrai que la mise en oeuvre de systèmes intelligents fait souvent intervenir des considérations d'ordre purement pratique reliées à l'environnement où sont implémentés ces systèmes, c'est néanmoins sur une connaissance solide des fondements théoriques en jeu que reposent de façon essentielle les progrès en IA. Un tel point de vue sur le rôle-clé de la logique mathématique en IA est soutenu avec force par de nombreux chercheurs, comme par exemple Genesereth et Nilsson qui fondent sur cette thèse leur ouvrage récent [GN87].

Si on peut faire remonter jusqu'à Aristote la recherche systématique de formes correctes de raisonnement (qu'on se rappelle ici la théorie scolastique du syllogisme dont est issue la célèbre litanie mnémotechnique "*Barbara Celarent Darii Ferio Cesare* (...)") rattachée aux 19 formes de syllogismes catégoriques), le véritable précurseur de la conception moderne de la logique fut sans doute Leibniz (1646-1716). L'idée de créer un langage formalisé doté de règles de déduction appropriées se trouve clairement exprimée dans son projet de construction d'une *langue caractéristique universelle* ("*lingua characteristica universalis*"), système symbolique permettant l'élaboration d'un *calcul du raisonnement* ("*calculus ratiocinator*") applicable à tout le champ de la pensée. C'est en prenant comme modèle la "*Methode des Mathematiciens*" (qui, dit-il, "n'a pas été suffisante pour découvrir tout ce qu'on pouvoit souhaiter d'eux [mais qui] a été au moins capable de les garantir des fautes" [Le77], p. 153) que Leibniz conçoit son dessein:

"[S]i l'on pouvoit trouver des caracteres ou signes propres à exprimer toutes nos pensées, aussi nettement et exactement que l'arithmetique exprime les nombres, ou que l'analyse geometrique exprime les lignes, on pourroit faire en toutes les matieres autant qu'elles sont sujettes au raisonnement tout ce qu'on peut faire en Arithmetique et en Geometrie.

Car toutes les recherches qui dependent du raisonnement se feroient par la transposition de ces caracteres, et par une espece de calcul; ce qui rendroit l'invention de belles choses tout a fait aisée. Car il ne faudroit pas se rompre la teste autant qu'on est obligé de faire aujourd'huy, et neantmoins on seroit asseuré de pouvoir faire tout ce qui seroit faisable.

De plus on feroit convenir tout le monde de ce qu'on auroit trouvé ou conclu puisqu'il seroit aisé de verifier le calcul soit en le refaisant, soit en essayant quelques preuves semblables à celle de l'abjection novenaire [preuve par neuf] en arithmetique. Et si quelqu'un doutoit de ce que j'aurois avancé, je luy dirois: contons, Monsieur, et ainsi prenant la plume et de l'encre, nous sortirions bientost d'affaire." [Le77], pp. 155-156.

Leibniz accordait une très grande importance à ce projet de définition d'un langage symbolique et d'un système de règles grâce auquel tout raisonnement serait ramené à un nombre fini de transformations formelles:

"J'ose dire que cecy est le dernier effort de l'esprit humain, et quand le projet sera executé, il ne tiendra qu'aux hommes d'estre heureux puisqu'ils auront un instrument qui ne servira pas moins à exalter la raison, que le Telescope ne sert à perfectionner la vuee.

C'est une de mes ambitions de venir à bout de ce projet si Dieu me donne la vie. (...) Et comme je suis asseuré qu'il n'y a point d'invention qui approche de celle cy, je croy qu'il n'y a rien de si capable d'eterniser le nom de l'inventeur." [Le77], p. 157.

Il a fallu attendre le milieu du XIX^e siècle pour voir le premier développement d'une logique symbolique telle que suggérée par Leibniz, d'abord avec les travaux de Boole puis ceux de Frege et de Russell. (Frege se percevait d'ailleurs lui-même clairement comme travaillant dans l'esprit de Leibniz; voir à ce sujet [vH67], p. 2.) Mais les résultats limitatifs obtenus, au cours de la décennie 1930, par Gödel, Church et Turing sont venus mettre un terme au projet formulé par Leibniz: la mécanisation complète du raisonnement est impossible. Même dans le cadre précis d'un système formalisé comme celui de la logique

du premier ordre, il n'est pas possible, comme nous l'apprennent les travaux de Church et Turing sur l'indécidabilité de l'*Entscheidungsproblem*, de déterminer de façon routinière et infallible la validité d'énoncés quelconques.

Malgré de telles limitations inhérentes à tout formalisme, les progrès marqués de la logique mathématique au cours de la première moitié du XXe siècle ont créé une situation telle que, suite à l'apparition des premiers ordinateurs au cours des années 50, s'est manifesté avec force l'intérêt pour l'élaboration de systèmes informatiques aptes à effectuer des raisonnements dans certains contextes formels. Une partie des travaux a alors porté de façon spécifique sur la question de la déduction automatisée (on trouvera dans [Da83] et [WH83] de précieuses indications historiques à ce propos) tandis que d'autres recherches ont examiné la question générale de la représentation de la connaissance et du raisonnement de sens commun. Dans son travail fondamental [Mc58], reconnu comme ayant démarré le sujet, McCarthy propose un système basé sur la logique du premier ordre où la connaissance nécessaire au programme est exprimée sous forme déclarative et où le raisonnement se fait par le biais d'opérations logiques sur les énoncés.

Les années 60 et surtout les années 70 ont été témoins d'une prolifération de travaux où diverses questions d'IA étaient examinées principalement dans le cadre de la logique des prédicats du premier ordre. Mais peu à peu divers inconvénients provoqués par un certain manque de souplesse de cette logique ont été relevés, ce qui a suscité un intérêt pour l'utilisation d'autres logiques. Dans certains cas, il s'est agi d'aller chercher des logiques étudiées depuis fort longtemps par logiciens et philosophes, alors que dans d'autres cas de nouveaux systèmes ont dû être créés de toutes pièces. Là encore ces recherches ont provoqué l'éclosion d'une littérature fort abondante. (On en retrouvera les textes fondamentaux dans les anthologies [WN81], [BL85] et [Gi87].)

Le but du présent travail est d'examiner le cadre théorique général de l'utilisation de la logique mathématique en IA, plus précisément en ce qui concerne le raisonnement automatisé, et de faire ressortir le contexte particulièrement approprié et fécond que détermine la logique. Après un rappel des rudiments de la logique du premier ordre, nous présentons de façon sommaire quelques systèmes logiques non classiques qui se retrouvent au coeur des travaux courants et qui ont été introduits en vue de corriger certaines déficiences de la logique classique du point de vue du raisonnement automatisé, en particulier en ce qui concerne le raisonnement relevant du sens commun. Cet exposé, qui se veut essentiellement non technique, vise donc à offrir une vue unifiée des principales avenues de recherche actuelles de façon à mettre en relief les grandes lignes de l'évolution récente du domaine. Le lecteur désireux de pousser plus à fond son étude trouvera en référence tant des articles de base dans lesquels furent introduits les systèmes dont il est question ici que des ouvrages récents traitant diverses facettes du sujet de façon plus détaillée. Le thème dont nous faisons l'exposition faisant actuellement l'objet d'une littérature extrêmement abondante, nous avons voulu ainsi fournir au lecteur certains jalons bibliographiques.

L'auteur se présente ni comme un juge, ni comme un expert ou un défenseur de l'un ou l'autre des systèmes dont il sera question ici. C'est plutôt en amateur (au sens noble du terme!) qu'il aborde ces questions et il souhaite qu'un tel point de vue puisse être profitable à tous ceux qui partagent avec lui ce statut.

2. Logique et logiques. Un système logique (ou plus brièvement une *logique*) peut être vu comme un système formel permettant de représenter la connaissance à l'aide d'énoncés formulés dans un langage approprié. Comme l'étude d'un tel système porte tant sur la manipulation formelle des énoncés que sur l'interprétation qu'on peut en faire, chaque système logique comporte ainsi trois aspects, l'un concernant la *représentation*, l'autre la *déduction formelle* et le dernier l'*interprétation*. Une logique est donc déterminée par la donnée d'une *syntaxe* permettant de produire des expressions formelles dans un certain langage, d'un *calcul déductif* opérant sur ces expressions en vue de tirer une conclusion à partir d'un ensemble de prémisses donné et d'une *sémantique* fournissant une interprétation aux expressions. (Une caractérisation générale, dans le contexte de la théorie des catégories, des diverses "composantes" d'un système logique a été introduite récemment dans [Me89].) Le choix d'une logique appropriée à la représentation de connaissances données est sujet à discussion: il n'existe pas de règles pour établir si le système choisi est le *bon* système, non plus qu'il est possible de déterminer de façon absolue si un certain ensemble de formules décrit bien une situation donnée.

2.1 Logique classique. Afin de mieux mesurer l'apport des logiques non classiques, il convient d'abord de rappeler brièvement certaines caractéristiques de la logique classique, c'est-à-dire de la *logique des prédicats du premier ordre* (le *calcul des propositions* étant à toutes fins utiles une "miniaturisation" de ce système). Tant en informatique qu'en IA, la logique classique a joué, et continue à jouer, un rôle de première importance qui demande qu'on la connaisse beaucoup plus intimement que comme simplement "la logique qui n'est pas non classique" [SM88], p. 1.

2.1.1 Syntaxe et sémantique. Le langage de la logique du premier ordre est défini à partir d'un alphabet comprenant, outre les symboles logiques usuels (connecteurs booléens, quantificateurs), des symboles de variables, de constantes, de fonctions et de prédicats. Ces symboles sont assemblés selon certaines règles précises (nous n'entrerons pas ici dans les détails techniques, qui se retrouvent dans tout livre d'introduction à la logique mathématique – voir par exemple [Me64] ou [vD83]) de façon à former deux catégories d'expressions formelles: les termes et les formules. Il s'agit là d'objets syntaxiques qui représentent respectivement des objets et des relations dans notre "univers du discours" et qui demandent à être interprétés à l'aide d'une structure sémantique appropriée. Nous touchons ainsi du doigt le double traitement que l'on peut faire de la logique, sous ses aspects sémantique et syntaxique.

Dans l'approche *sémantique*, une structure est introduite qui sert de lieu pour interpréter les expressions du langage. On obtient ainsi la notion de formule *vraie* (pour l'interprétation en cause – il y aura en général différentes façons d'interpréter une formule donnée dans différents univers). Une formule vraie dans toute interprétation est dite *valide*. Désignant par *théorie* un ensemble de formules du langage, on peut alors formaliser l'idée qu'une formule découle de certaines autres: la formule α est une *conséquence (sémantique)* de la théorie Δ (ce que l'on écrit $\Delta \models \alpha$) si elle est vraie dans toute interprétation qui rend les formules de Δ vraies (c'est-à-dire dans tout *modèle* de la théorie Δ).

L'approche *syntactique*, quant à elle, fait voir les symboles du langage comme de véritables "taches d'encre" (ou plus précisément des classes d'équivalence de taches d'encre – voir [Ha82], p. 16) que l'on peut manipuler de façon purement formelle en suivant certaines règles combinatoires, tout comme dans le projet de Leibniz: on n'a pas alors à faire intervenir ce que pourrait être la signification d'une formule donnée. Une *dérivation formelle* consiste ainsi en une suite finie de transformations selon des règles d'inférence bien choisies; une telle manipulation est donc sujette à être éventuellement automatisée. Lorsqu'il est

possible de dériver formellement une formule α à partir d'une théorie Δ (selon un certain calcul déductif), alors α est dite *conséquence (syntaxique)* de Δ (ce que l'on écrit $\Delta \vdash \alpha$). Lorsque $\Delta = \emptyset$, on dit que α est un *théorème*.

On reconnaît traditionnellement trois principaux types de techniques déductives qui ont été développées pour effectuer des dérivations formelles (voir [Su83] pour une synthèse particulièrement limpide). Dans un *système hilbertien*, certaines formules sont identifiées comme axiomes et l'on introduit quelques règles d'inférence, de façon typique le Modus Ponens

$$\frac{\alpha \quad \alpha \supset \beta}{\beta}$$

(lire: "de α et de $\alpha \supset \beta$, on peut déduire β "; ici, $\supset$ est le symbole de l'implication matérielle) et une règle pour les quantificateurs. (Une règle d'inférence peut être vue en général comme un $(n+1)$ -uplet $\langle E_1, \dots, E_n, E \rangle$ dont la signification est que l'expression formelle E peut être déduite des expressions $E_1, \dots, E_n$.) Un *système de déduction naturelle* (notion due à Gentzen) remplace les axiomes par une série de règles régissant l'introduction ou l'élimination des connecteurs et quantificateurs. Par exemple, pour la conjonction $\wedge$, on aura les deux règles suivantes:

$$\frac{\alpha \quad \beta}{\alpha \wedge \beta} \qquad \frac{\alpha \wedge \beta}{\alpha}$$

Gentzen a également introduit un *calcul des séquents* qui généralise l'approche précédente en "internalisant" dans le système formel la méta-opération de dérivation. Un séquent est une expression formelle du type $\alpha_1, \dots, \alpha_n \rightarrow \beta_1, \dots, \beta_m$ dont le sens est que l'un des β_j est vrai si tous les α_i le sont: les "hypothèses" α_i permettent donc d'atteindre l'un des "buts" β_j . (Les définitions précises de chacun de ces trois systèmes d'inférence se retrouvent facilement dans la littérature, entre autres dans les manuels [Me64], [vD83] et [Ga86] respectivement.)

2.1.2 Le théorème de complétude. Un théorème fondamental, dû à Gödel (1930), nous renseigne sur le lien entre la notion syntaxique de $\vdash$ et la notion sémantique de $\models$: c'est le *théorème de complétude* qui énonce l'équivalence

$$\Delta \vdash \alpha \iff \Delta \models \alpha.$$

(En particulier, lorsque Δ est vide, ce résultat exprime l'identité des théorèmes avec les formules valides.) Il est donc possible de formuler des systèmes déductifs qui sont non seulement *sains* ($\implies$) (c'est bien sûr une propriété minimale d'un "bon" système déductif que n'importe quelle formule déduite d'une théorie Δ soit "validée" par tout modèle de Δ : la déduction formelle ne doit pas produire d'objet incongru), mais également *complets* (pour la déduction) ($\impliedby$) en ce sens que toute conséquence sémantique de Δ peut être déduite formellement à partir de Δ . Ce résultat est fondamental pour la question du raisonnement automatisé parce qu'il suggère de remplacer l'examen de *toutes* les situations possibles où l'ensemble de prémisses Δ est réalisé (vision sémantique) par une étude de nature combinatoire des formules de Δ (vision syntaxique) se prêtant davantage à l'automatisation. Une telle approche à la validité a toutefois ses limitations: il n'est pas possible, en vertu des *théorèmes d'indécidabilité* de Church et de Turing (1936), de construire une procédure générale permettant de distinguer les théorèmes des non-théorèmes en logique du premier ordre.

Il est intéressant de remarquer que dans le cas plus restreint du calcul des propositions, le problème de la validité a une solution effective: on peut déterminer si une formule quelconque est une *tautologie* simplement en examinant toutes les interprétations correspondantes (en nombre fini!). Mais une telle procédure n'est pas "efficace" (au sens de la théorie de la complexité du calcul) puisqu'elle requiert un temps exponentiel en fonction de la longueur de la formule. Tout semble d'ailleurs indiquer, eu égard aux résultats sur la $\mathcal{NP}$ -complétude, que ce problème ne peut être résolu en temps raisonnable (voir à ce propos les commentaires de [Ga86], p. 54).

2.1.3 La résolution. Les différentes familles de systèmes déductifs mentionnées précédemment présentent chacune leurs particularités. Les systèmes hilbertiens reposent sur un petit nombre de règles d'inférence, ce qui rend leur étude théorique particulièrement limpide; toutefois ils font appel à des axiomes quelque peu parachutés et l'exécution de véritables dérivations y est réellement laborieuse et pas du tout intuitive. C'est justement pour mieux faire ressortir l'intuition derrière les preuves formelles que Gentzen a conçu la déduction naturelle, qui aspire à reproduire le plus fidèlement possible la façon de faire du mathématicien. Le calcul des séquents fournit quant à lui une approche plus facile à exprimer sous forme algorithmique (c'est là la thèse défendue dans [Ga86]); il est intimement lié à la technique des *tableaux* qui semble connaître récemment un regain d'intérêt (voir par exemple [dS88]). Toutefois aucun de ces systèmes n'avait été conçu à l'origine en vue d'une véritable automatisation du raisonnement: ils sont "human-oriented" plutôt que "machine-oriented", pour reprendre les termes de Robinson [Ro65]. Il a fallu attendre les années 60 pour voir les premiers développements de techniques, fondées sur le *théorème de Herbrand* (1930), cherchant à tirer profit des possibilités de l'ordinateur.

Mais c'est surtout grâce à l'introduction du *principe de résolution* par Robinson [Ro65] que s'est ouverte la voie permettant toute une série de recherches fructueuses sur la déduction automatisée en logique classique (on consultera à ce sujet les exposés [Da83], [WH83] et [WV87]). L'idée à la base de la résolution est assez naturelle, cette règle étant une généralisation du Modus Ponens, comme on le voit aisément dans le cadre propositionnel:

$$\frac{\alpha \vee \gamma \quad \beta \vee \neg \gamma}{\alpha \vee \beta}$$

(on n'a qu'à considérer le cas où α est faux). La résolution a un bon comportement théorique. On vérifie d'abord aisément qu'il s'agit d'une règle saine; et bien qu'elle ne soit pas complète selon le sens mentionné plus haut (ainsi la tautologie $\alpha \vee \neg \alpha$ ne peut être obtenue par résolution à partir de la théorie vide), elle jouit néanmoins d'une propriété de complétude: elle est *complète pour la réfutation* en ce sens qu'à partir d'une théorie inconsistante (c'est-à-dire sans modèle), la résolution produit toujours une contradiction. (On trouvera plus de détails sur ces questions dans [GN87].)

Sur le plan pratique, le principe de résolution a donné lieu au développement de stratégies variées afin d'en améliorer l'efficacité en cherchant à mieux contrôler la complexité de l'espace de recherche; encore aujourd'hui la résolution demeure au centre de nombreuses recherches (dont on trouvera écho dans [WV87]). En particulier, la résolution joue un rôle-clé dans le contexte de la programmation logique qui, via un langage tel PROLOG, fournit des implémentations particulièrement fécondes de (fragments substantiels de) la logique classique. (Une présentation générale de la logique de forme clausale, qui constitue le cadre théorique de PROLOG, est l'objet du livre [Ri89]; les liens entre PROLOG et les principes de résolution et d'unification sont également examinés dans [Ga86] et [Th8-].)

2.2 Logiques non classiques.

2.2.1 Logique modale. Avant de parler du rôle des logiques non classiques en IA, il est utile d'indiquer les origines de certaines d'entre elles. L'idée à la base de l'introduction de nouvelles logiques est qu'il est parfois souhaitable de rendre plus souples les moyens d'expression dont nous disposons en logique du premier ordre. L'un des exemples les plus anciens de telles modifications se retrouve dans la *logique modale*; cette expression recouvre une large famille de systèmes logiques formalisant des "opérateurs de modalité" qui viennent modifier la portée des énoncés. Déjà la logique aristotélicienne s'intéressait à une telle approche et distinguait quatre modalités s'opposant deux à deux: le nécessaire et le contingent, le possible et l'impossible. Ces façons de qualifier les énoncés sont connues sous le nom de modalités *aléthiques* puisqu'elles concernent la notion de vérité. En symboles, on écrit $\Box \alpha$ pour "il est nécessaire que α " ou "nécessairement α " et $\Diamond \alpha$ pour "il est possible que α " ou "éventuellement α ", ce qui conduit à des équivalences telle $\Box \alpha \iff \neg \Diamond \neg \alpha$. Selon le genre de nuances que l'on a à l'esprit, on peut introduire de nombreux autres opérateurs modaux: modalités *épistémiques*, concernant la connaissance: " X sait que α ", " X croit que α "; modalités *déontiques*, concernant les devoirs: "il est interdit que ...", "il est obligatoire que ..."; modalités *temporelles*: "il est toujours vrai que ...", "il est parfois vrai que ..."; etc.

Contrairement à ce qui se passe en logique classique, l'effet d'un opérateur modal ne dépend pas strictement de la valeur de vérité de l'énoncé auquel il est appliqué. Considérons par exemple le cas de l'opérateur $\Box$. Certaines affirmations représentent des vérités absolues et immuables alors que d'autres sont tantôt vraies et tantôt fausses selon le point de vue, le moment, etc. On est ainsi amené à distinguer entre des "niveaux de vrai", dont certains sont plus contraignants que d'autres: certaines vérités sont des "vérités nécessaires" et d'autres sont des "vérités contingentes". Ainsi si α est faux, alors il en est clairement de même pour $\Box \alpha$; mais si α est vrai, la "valeur de vérité" de $\Box \alpha$ n'est pas déterminée. Par exemple il est nécessairement vrai que $2 + 2 = 4$; mais même s'il fait beau actuellement, cela n'est pas une vérité nécessaire: le mauvais temps demeure toujours possible.

Les travaux modernes en logique modale, initiés principalement par Lewis durant la première partie de ce siècle, ont amené l'élaboration d'une kyrielle de systèmes axiomatiques visant à préciser l'intuition derrière les opérateurs modaux. Si le statut de certains axiomes, par exemple $\Box(\alpha \supset \beta) \supset (\Box \alpha \supset \Box \beta)$, ne semblait pas poser problème, il en allait tout autrement pour des énoncés tels $\Box \alpha \supset \Box \Box \alpha$ ou $\Diamond \alpha \supset \Box \Diamond \alpha$. On a ainsi assisté à l'identification de plusieurs systèmes déductifs différents, les axiomes de base variant de l'un à l'autre. Une littérature abondante s'est développée autour des propriétés déductives de ces systèmes dont les plus célèbres sont désignés par T , $S4$ et $S5$ (voir [HC68] pour une exposition détaillée). Il a fallu attendre l'élaboration de la *sémantique des mondes possibles* de Kripke [Kr63] pour avoir une analyse sémantique rigoureuse venant clarifier le sens de $\Box$ et de $\Diamond$. L'idée de Kripke a été de multiplier les possibilités d'interprétation d'une formule modale; il considère donc un "univers" d'interprétation constitué de différents "mondes possibles" qui, intuitivement, correspondent à diverses façons dont les choses peuvent se produire. Ces mondes sont reliés entre eux par une relation d'"accessibilité" et Kripke a démontré que les différences entre les divers systèmes axiomatiques pour la logique modale correspondent à certaines restrictions sur cette relation d'accessibilité. Dans une structure de Kripke, c'est-à-dire dans un univers d'interprétation donné, la valeur de vérité d'une formule n'est plus absolue mais dépend du monde possible considéré. En particulier, l'idée de nécessité est rendue par le fait qu'on stipule que l'énoncé $\Box \alpha$ est vrai dans un monde donné s'il est vrai dans *tous* les mondes accessibles à partir de ce monde. De même $\Diamond \alpha$

est vrai dans un monde *s'il existe* un monde accessible où il est vrai. On remarquera ici qu'une telle interprétation des opérateurs modaux n'est pas sans rappeler le rôle des quantificateurs en logique du premier ordre: il est en effet possible de "simuler" des énoncés de logique modale en logique du premier ordre (voir [GN87]). Mais comme on y perd alors en clarté et en concision, la logique modale conserve toujours son utilité et sa raison d'être. (On trouvera un exposé succinct des principes de la logique modale, surtout en ce qui concerne son lien avec l'IA, dans [SM88], [Th8-] et [ID90].)

2.2.2 Autres logiques non classiques. Outre la logique modale, plusieurs autres logiques non classiques ont été introduites afin de relâcher certaines contraintes de la logique du premier ordre. Ainsi les *logiques multivalentes* visent à sortir de la dichotomie du vrai et du faux par suite de la constatation suivante: il n'est pas toujours possible d'attribuer une valeur de vérité précise, vrai ou faux, à chaque affirmation, que ce soit par manque d'information (" $x = 7$ ": que représente x ?) ou par manque d'un critère précis ("il fait beau": qu'est-ce que le beau temps?). On a cherché à modéliser ces nuances en introduisant des logiques basées sur plus de deux valeurs de vérité. Déjà au niveau d'une logique *trivalente*, dans laquelle intervient une troisième valeur de vérité X , une telle approche donne lieu à plusieurs systèmes différents selon l'intuition que l'on cherche à représenter. Les calculs trivalents sont conçus comme des extensions du calcul propositionnel classique mais ils se distinguent les uns des autres par la façon dont ils opèrent sur X . Ainsi le tableau suivant présente trois définitions possibles, dues respectivement à Lukasiewicz (I), Bochvar (II) et Kleene (III), des valeurs de vérité de l'énoncé $\alpha \supset \beta$ lorsque α ou β ont comme valeur X . (Dans chacun de ces trois calculs, la valeur de vérité de $\neg \alpha$ est X lorsque celle de α est X .) L'intuition derrière la convention de Bochvar, par exemple, est que X représente un énoncé sans signification, non pertinent. Les conséquences de ces définitions sont parfois remarquables. Ainsi l'équivalence $\alpha \supset \beta \iff \neg \alpha \vee \beta$ est valide dans le système de Kleene, mais pas dans celui de Lukasiewicz; par contre, $\alpha \supset \alpha$ est un tautologie pour Lukasiewicz, mais pas pour Kleene. (Pour plus de détails sur ces systèmes ainsi que sur leur développement historique, on consultera [Re68], [Tu84] ou [ID90].)

α	β	$\alpha \supset \beta$ (I)	$\alpha \supset \beta$ (II)	$\alpha \supset \beta$ (III)
V	X	X	X	X
F	X	V	X	V
X	V	V	X	V
X	F	X	X	X
X	X	V	X	X

De façon analogue, on peut mettre en place des logiques multivalentes comportant un nombre fini ou même infini de valeurs de vérité. Les premières constructions de ces systèmes sont dues à Lukasiewicz au cours des années 1920–30 et ont préparé le terrain tant pour une vision *probabiliste* de la logique, où chaque énoncé reçoit un "coefficient de vraisemblance" (entre 0 et 1), que pour l'utilisation de la notion de *flou* en logique. Nous verrons dans la prochaine section comment de telles approches ont fourni des outils nouveaux pour aborder certains problèmes reliés au raisonnement automatisé.

Lorsqu'on s'intéresse à des logiques différentes de la logique du premier ordre, il y a plusieurs autres systèmes qui viennent à l'esprit. On peut penser par exemple à l'*intuitionisme*, qui, en restreignant l'existence mathématique à une existence "constructive", a joué un rôle fort important sur le plan des fondements des mathématiques. Ou encore aux *logiques d'ordre supérieur*, qui viennent enrichir les moyens d'expression de la

logique du premier ordre en prolongeant la quantification aux symboles de prédicats. Ou même aux *logiques infinitaires*, qui permettent l'écriture de conjonctions ou de disjonctions de longueur infinie. Il ne sera pas question de tels systèmes dans la suite de ce texte. Soulignons cependant qu'outre sa portée philosophique, l'intuitionisme a exercé une influence en informatique théorique, comme en témoignent les travaux de Martin-Löf sur les fondements de la théorie de la programmation (voir [Tu84]) et ceux de Gabbay sur le raisonnement non classique (voir [SM88]). Quant aux logiques infinitaires ou d'ordre supérieur, les recherches récentes ont surtout porté sur le développement d'une "théorie abstraite des modèles" visant à fournir un contexte unificateur pour comparer les propriétés de divers systèmes (on consultera à ce sujet [BF85]) et semblent avoir eu une influence restreinte en IA. Bibel [Bi86] se fait toutefois l'avocat d'une étude approfondie des logiques d'ordre supérieur: comme elles sont trop riches et trop complexes dans toute leur généralité, il s'agit donc d'en identifier de "bons" fragments susceptibles d'être plus propices pour l'IA.

Si l'étude de logiques autres que la logique du premier ordre est depuis longtemps un sujet d'intérêt en logique mathématique (on trouvera dans l'ouvrage de base [Ba77] des exposés sur l'intuitionisme, les logiques d'ordre supérieur et infinitaires), il est cependant intéressant de constater que les systèmes logiques non classiques qui semblent avoir exercé le plus d'influence en informatique, à tout le moins en IA, et dont il sera question dans la prochaine section, avaient souvent été perdus de vue par les mathématiciens au cours des dernières années. Ainsi les logiques modales et multivalentes sont complètement absentes de [Ba77], bien qu'elles aient été étudiées par de nombreux logiciens mathématiciens plus tôt au cours de ce siècle, comme il appert des remarques historiques de [Re68]. C'est plutôt dans des ouvrages de logique philosophique, tel [GG8-], que se retrouve habituellement la présentation de systèmes de cette nature. Mais les besoins récents de l'informatique, et en particulier de l'IA, ont remis ces logiques non classiques parmi les sujets d'intérêt actuel en logique mathématique.

3. Inférence dans un contexte de connaissances imparfaites. Certains systèmes d'IA ont été conçus pour effectuer des raisonnements dans des contextes où la connaissance de l'"univers" est complète. Ainsi un programme joueur d'échecs sait la position de toutes les pièces sur l'échiquier et un démonstrateur automatique connaît tous les axiomes de la théorie formelle où il doit oeuvrer. Mais d'autres systèmes, en vue de mieux cerner l'idée de raisonnement effectué selon le "sens commun", tenteront d'effectuer des déductions dans des contextes plus près de la "vraie vie", où les connaissances sont souvent imparfaites: connaissance partielle de certains faits, absence de faits essentiels, voire contradiction entre des faits connus. La logique classique, malgré son pouvoir expressif et ses capacités déductives, n'est pas vraiment appropriée pour de telles situations, puisqu'elle s'attache à la formalisation du raisonnement strictement correct et ne sait rendre compte de raisonnements qui, bien que judicieux, comporteraient une certaine part d'incertitude. Zadeh [Za84] décrit la chose en ces termes: "The crux of the problem, really, is the excessively wide gap between the precision of classical logic and the imprecision of the real world."

Il est vrai pourtant que plusieurs concepts de base de la logique du premier ordre concernent précisément la description de situations de connaissances imparfaites, comme le souligne [Mo82]: ainsi la disjonction de deux formules nous permet de dire qu'au moins l'une d'entre elles est vraie sans savoir laquelle; la quantification existentielle nous permet d'affirmer l'existence d'un objet ayant une certaine propriété sans connaître cet objet; et la quantification universelle nous permet de dire que toute une classe d'objets possède une

propriété sans connaître explicitement chacun des objets de cette classe. (Cette observation à propos de la logique du premier ordre est illustrée dans [Mo82] à l'aide du célèbre exemple des trois cubes: trois cubes, A , B , et C , sont placés dans l'ordre à la queue leu leu; on sait que A est vert et que C est bleu, la couleur de B étant inconnue; on demande s'il y a un cube vert contigu à un cube qui n'est pas vert.) Mais les propriétés requises des systèmes logiques en IA quant au traitement de connaissances imparfaites vont au delà de telles caractéristiques: il s'agit, face à une information incertaine, incomplète ou évolutive, de pouvoir effectuer des raisonnements éventuellement sujets à révision. Les formalismes dont il sera question dans cette section représentent diverses réponses devant la nécessité d'un tel raisonnement révisable.

3.1 Logiques non monotones.

3.1.1 Monotonie de la logique classique. On a déjà mentionné à propos de certains des systèmes déductifs considérés plus haut qu'il sont sains, ce qui est une caractéristique fort rassurante: on souhaite bien sûr que les formules obtenues par manipulation formelle soient interprétées comme étant vraies. Mais en un certain sens, cette propriété d'intégrité fait que la dérivation formelle ne produit pas de connaissance vraiment "nouvelle" à propos de l'univers sous étude. En effet, supposons que l'information disponible ait été colligée en une "base de connaissances", c'est-à-dire un ensemble de formules Δ , et que l'on ait sous considération certains modèles de la théorie Δ , donc certaines interprétations rendant vraies toutes les formules de Δ . Si α est une formule telle que $\Delta \vdash \alpha$, alors, comme $\vdash$ est sain, α est vraie dans tout modèle de Δ . Le fait d'avoir dérivé α n'a donc pas permis d'affiner notre connaissance de l'univers, en ce sens que les modèles de $\Delta \cup \{\alpha\}$ ne sont autres que ceux de Δ . Bien sûr, la dérivation de α a toutefois rendu explicite une information qui était en quelque sorte contenue implicitement dans Δ (c'est justement là le propre d'un système sain).

De ce point de vue, l'enrichissement de notre connaissance à propos de l'univers ne se fera pas tant par la dérivation de tout ce que notre base de connaissances contient potentiellement, que par l'ajout d'informations nouvelles ou encore par la révision de certaines informations déjà connues. Or la logique du premier ordre comme telle ne nous fournit pas d'outils permettant de modifier notre base de connaissances. En effet, la relation de dérivabilité en logique classique jouit d'une propriété fondamentale de *monotonie*:

$$\Delta \vdash \alpha \implies \Delta \cup \{\beta\} \vdash \alpha.$$

L'ensemble des formules dérivables croît donc avec la base Δ . Si ceci est raisonnable dans le cas où Δ correspond à un ensemble d'hypothèses pour une théorie axiomatique (plus on a de données, plus on peut en tirer de résultats), ce ne l'est plus si on veut utiliser Δ comme une banque d'informations dynamique sur une situation donnée. Si la nouvelle information β vient à l'encontre de α , la logique classique n'a qu'une constatation possible: CONTRADICTION! et la nouvelle base $\Delta \cup \{\beta\}$ devient par le fait même totalement inintéressante. L'idée à l'origine des *logiques non monotones*, dont les premiers systèmes remontent à la fin des années 70, est de récupérer une telle situation en permettant la révision de nos bases de connaissances par l'ajout d'informations perçues comme plus fiables. Contrairement à ce qui se passe en logique classique, une structure d'inférence non monotone n'est pas valide: les conclusions inférées ne sont pas nécessairement vérifiées dans tous les modèles des prémisses et ne sont acceptées que jusqu'à preuve éventuelle du contraire. Ainsi le constat de contradiction de la logique classique est évité.

La non monotonie correspondant à un aspect fondamental du raisonnement de sens commun, il n'est donc pas surprenant de la retrouver au coeur de nombreux formalismes

pour la représentation des connaissances. Cette même idée est également présente dans des contextes plus restreints, tels celui de la *négation par échec* [CI78] ou de l'*hypothèse du monde clos* [Re78]. Nous ne pouvons, dans le cadre de ce texte, qu'évoquer les grandes lignes des approches non monotones en IA. On trouvera plus de détails dans de nombreux exposés récents, entre autres dans [Et88], [GN87], [Re87], [SM88], [So89] ou [Th8-]; on pourra aussi consulter l'anthologie [Gi87] de même que le numéro spécial de la revue *Artificial Intelligence* (13 (1980) pp. 1-174) dans lequel ont paru les articles de base [Mc80], [MD80] et [Re80], dont nous esquissons maintenant les contributions.

3.1.2 Exemples de systèmes non monotones. La logique non monotone de McDermott et Doyle [MD80] prolonge la logique du premier ordre par l'adjonction d'un opérateur modal $\mathcal{M}$ qui correspond intuitivement à l'idée de cohérence (déductive). La formule $\mathcal{M}\alpha$ signifie que α est cohérente avec ce que l'on croit, en ce sens que $\neg\alpha$ ne peut être dérivée de la base de connaissances actuelle. C'est de façon typique par l'introduction d'axiomes de la forme $(\beta \wedge \mathcal{M}\alpha) \supset \alpha$ que se manifeste l'effet de la modalité $\mathcal{M}$. McDermott et Doyle donnent comme exemple le système d'axiomes suivants, où μ signifie "il est midi", σ "il fait soleil" et ε "il y a une éclipse".

- (1) $(\mu \wedge \mathcal{M}\sigma) \supset \sigma$,
- (2) μ ,
- (3) $\varepsilon \supset \neg\sigma$.

Comme cette base de connaissances ne permet pas de dériver $\neg\sigma$, $\mathcal{M}\sigma$ est interprétée comme vraie, d'où on en tire σ par (1) et (2). Si maintenant on ajoute l'information

- (4) ε ,

alors σ ne peut plus être dérivée: il y a donc eu révision du raisonnement antérieur.

La modalité $\mathcal{M}$ correspond ainsi à une espèce de méta-règle affirmant:

$$\not\vdash \neg\alpha \implies \vdash \mathcal{M}\alpha.$$

Mais comme elle fait intervenir la *non dérivabilité* d'une formule dans un contexte donné, une telle règle ne peut bien sûr être appliquée de façon effective. Le problème de définir une sémantique appropriée pour la dérivation non monotone amène McDermott et Doyle à introduire la méthode (non constructive) des *points fixes*: étant donné un ensemble de formules A , on s'intéresse aux solutions de l'équation

$$T = Th(A \cup \{\mathcal{M}\alpha \mid \neg\alpha \notin T\})$$

où Th représente l'opération habituelle de fermeture par rapport à la dérivation en logique du premier ordre. La notion de point fixe vise ainsi à capter l'idée que si $\neg\alpha$ n'est pas dérivable, alors $\mathcal{M}\alpha$ l'est. La logique non monotone présentée dans [MD80] comporte certaines difficultés, comme les auteurs le soulignent eux-mêmes: ainsi $\{\mathcal{M}\alpha, \neg\alpha\}$ n'est pas un ensemble contradictoire tandis que $\{\mathcal{M}\alpha \supset \beta, \neg\beta\}$ l'est (puisque $\neg\alpha$ ne peut être déduite, alors on a $\mathcal{M}\alpha$); de même $\mathcal{M}\alpha$ ne découle pas de $\mathcal{M}(\alpha \wedge \beta)$. C'est en vue de pallier à de telles lacunes que McDermott [Mc82] a introduit d'autres logiques basées sur les systèmes modaux T , $S4$ et $S5$ mentionnés précédemment. On trouvera dans [Mo88], [Re87] et [So89] des commentaires sur la portée et les limitations de ces diverses variantes.

En réponse aux difficultés de l'approche de McDermott et Doyle, Moore a proposé une logique dite *autoépistémique* où le concept de cohérence est remplacé par celui de croyance. Ce système met en scène un agent doté de certaines capacités introspectives par rapport à ce qu'il croit. Moore utilise à cette fin un opérateur de modalité épistémique $\mathcal{B}$ (pour

“believe”), la formule $B\alpha$ étant vérifiée si et seulement si α appartient à l’“ensemble des croyances” C de l’agent en cause. On considère certaines propriétés de stabilité sur C , comme par exemple

$$\begin{aligned}\alpha \in C &\implies B\alpha \in C & (“\alpha \text{ est crue}” \text{ est une croyance si } \alpha \text{ est une croyance}) \\ \alpha \notin C &\implies \neg B\alpha \in C & (“\alpha \text{ n'est pas crue}” \text{ est une croyance si } \alpha \text{ ne l'est pas}).\end{aligned}$$

La non monotonie résulte du fait que l’ensemble des croyances d’un agent peut se modifier avec le temps. [Mo88] présente un exposé détaillé de cette approche; on y trouve en particulier une interprétation sémantique en termes de mondes possibles: $B\alpha$ est vraie dans un monde donné si α est vraie dans tout monde accessible à partir de ce monde.

Une autre approche importante à la non monotonie est la *logique des défauts* de Reiter [Re80], dont l’objectif général est la formalisation d’inférences reposant sur des lois générales quoique susceptibles d’admettre des exceptions (on trouvera une présentation détaillée de la logique des défauts dans l’ouvrage [Be89]). Comme dans [MD80], cette approche utilise un mode de raisonnement par défaut: en l’absence d’informations à l’effet contraire, on peut accepter certains faits. Dans la logique de Reiter, ce principe est mis en oeuvre par le biais de règles d’inférence de la forme

$$\frac{\alpha \quad : \quad \beta}{\gamma}$$

qui sont intégrées au système, plutôt qu’à l’aide de formules spéciales du langage. L’interprétation d’une telle *règle de défaut* est la suivante: inférer γ si α est connue et si β est cohérente avec ce qui est connu (c’est-à-dire si $\neg\beta$ ne peut être dérivée). Comme l’inférence par défaut est ainsi définie en termes d’inférence, c’est par une technique de points fixes, tout comme dans la logique de McDermott et Doyle, que Reiter décrit le comportement de son système. Il introduit ainsi des extensions d’ensembles de formules représentant des états cohérents de l’“univers”. Pour Reiter, le raisonnement se fera en choisissant une extension et en y travaillant jusqu’à ce qu’une nouvelle information vienne modifier la situation et forcer le choix d’une autre extension.

Une *théorie avec défauts* est donc déterminée par la donnée d’un ensemble d’axiomes ainsi que d’un ensemble de règles de défaut. Pour mieux illustrer ces idées, nous reprenons ici l’archétype des exemples cités sur le sujet. Si l’affirmation

$$\forall x(\text{OISEAU}(x) \supset \text{VOLE}(x))$$

semble raisonnable à prime abord, on réalise vite que cette loi souffre de nombreuses exceptions: les oisillons ne volent pas, pas plus que les autruches par exemple. Bien que la logique classique permette d’exprimer de telles exceptions par le biais de l’implication

$$\forall x(\text{OISEAU}(x) \wedge \neg \text{OISILLON}(x) \wedge \neg \text{AUTRUCHE}(x) \supset \text{VOLE}(x)),$$

il ne s’agit pas là d’une approche heureuse puisque l’apparition de toute nouvelle exception nécessiterait la modification de la formule. De plus, on ne pourra conclure à la capacité de voler que des seuls oiseaux dont on sait qu’ils ne font pas partie des exceptions. Il est cependant commode de pouvoir admettre qu’un oiseau vole tant que le contraire n’a pas été affirmé: c’est précisément cet objectif que remplit la règle de défaut

$$\frac{\text{OISEAU}(x) \quad : \quad \text{VOLE}(x)}{\text{VOLE}(x)}.$$

L'introduction des exceptions précédentes dans le système se ramène donc à l'ajout des formules

$$\begin{aligned} \forall x(\text{OISILLON}(x) \supset \text{OISEAU}(x) \wedge \neg \text{VOLE}(x)), \\ \forall x(\text{AUTRUCHE}(x) \supset \text{OISEAU}(x) \wedge \neg \text{VOLE}(x)). \end{aligned}$$

Dans [Mc80], McCarthy introduit une autre approche au raisonnement non monotone basée sur le principe de *circonscription*. L'idée est de remplacer l'examen de toutes les extensions cohérentes de l'état actuel des connaissances par l'examen d'extensions *minimales*. La circonscription ressort donc à la même philosophie que le principe de parcimonie de Guillaume d'Occam (XIV^e siècle) (principe familièrement connu sous le nom de *rasoir d'Occam*), à savoir qu'on ne postule pas l'existence de choses sauf nécessité: étant donné une certaine base de connaissances Δ , les seuls individus dont on supposera qu'ils possèdent une certaine propriété P sont ceux pour lesquels P peut être établie à partir de Δ . On "circonscrit" ainsi l'extension du prédicat P en restreignant l'ensemble des individus pertinents: il n'y aura pas d'individu "exceptionnel" satisfaisant P tel que cela ne puisse être connu à partir de Δ . Formellement, McCarthy définit un *schéma d'axiome de circonscription* tel que, pour un Δ et un P donnés et pour tout individu x , si $\Delta \not\models P(x)$, alors $\text{Circon}(\Delta, P) \vdash \neg P(x)$, où $\text{Circon}(\Delta, P)$ représente l'ensemble Δ auquel on a ajouté le dit axiome de circonscription (pour P et Δ). Ainsi, pour traiter l'exemple canonique sur les oiseaux introduit plus haut, McCarthy effectue la circonscription du prédicat $\text{ANORMAL}(x)$ qui cerne l'intuition qu'un oiseau normal vole (on trouvera plus de détails sur cet exemple dans [Th8-]). La non monotonie du système résulte de l'axiome même de circonscription, comme l'illustre l'exemple suivant de McCarthy: si Δ dit que A , B et C sont des blocs, alors, par circonscription, ce sont les seuls blocs et en particulier un autre objet D n'est pas un bloc; mais si on ajoute à Δ l'information que D est aussi un bloc, il faut alors réviser la conclusion précédente.

3.1.3 Une approche comparative. Les détails techniques concernant les logiques non monotones sont souvent assez complexes et présentés dans des formalismes variés, ce qui rend leur étude comparative difficile. C'est dans cette optique qu'ont été introduites certaines théories unificatrices, telle l'approche sémantique de Shoham ([Sh87], [Sh88]). Shoham considère une structure d'ordre partiel strict $<$ sur les interprétations d'une théorie et introduit la notion de modèle préféré: une interprétation est un *modèle préféré* d'une formule α si cette interprétation est un modèle de α et s'il n'y a pas d'autre interprétation la précédant dans la relation de préférence $<$ qui soit un modèle de α . On obtient une notion de conséquence sémantique *préférentielle* $\models_{<}$ plus restrictive que la notion usuelle et qui se lit: $\Delta \models_{<} \alpha$ si α est vraie dans tout modèle préféré de la théorie Δ . La non monotonie résulte du fait que ce ne sont plus tous les modèles de Δ qui entrent en jeu comme dans le cas de $\models$, mais seulement une partie de ces modèles. En effet, dans le cas classique, de $\Delta \models \alpha$ suit $\Delta \cup \{\beta\} \models \alpha$ puisque tous les modèles de $\Delta \cup \{\beta\}$ sont forcément modèles de Δ . Mais $\Delta \cup \{\beta\}$ peut avoir des modèles préférés qui ne sont pas des modèles préférés de Δ , de sorte qu'on peut avoir $\Delta \models_{<} \alpha$ sans avoir nécessairement $\Delta \cup \{\beta\} \models_{<} \alpha$. Shoham utilise la notion de conséquence préférentielle pour donner une caractérisation de la monotonie: une logique préférentielle (c'est-à-dire avec $\models_{<}$ défini comme ci-haut) est monotone si et seulement si, pour toute formule α , β et γ , $\alpha \models_{<} \beta \supset \gamma$ entraîne $\alpha \wedge \beta \models_{<} \gamma$.

L'intérêt de la logique préférentielle réside dans la liberté quant au choix de la relation de préférence $<$. Shoham démontre comment il est possible de récupérer les différentes logiques non monotones mentionnées ci-haut en faisant varier le critère de préférence;

une tel contexte global facilite bien sûr la comparaison entre ces systèmes. Shoham se sert également de sa notion de préférence pour introduire une méthode de représentation d'information *temporelle* lui permettant d'effectuer des raisonnements dans des contextes où des changements d'états interviennent.

3.1.4 Modalité et raisonnement révisable. On aura remarqué que le concept de modalité intervient de façon explicite dans certaines des logiques non monotones présentées plus haut. Ceci suggère de faire une étude plus approfondie sur le rôle que la logique modale peut jouer dans la représentation des connaissances et l'automatisation du raisonnement de sens commun. De nombreux travaux ont été consacrés à cette question et on en trouvera écho par exemple dans [GN87], [SM88], [ID90] ou [Wa90]. On peut également souligner à cet égard les résultats obtenus par Fariñas del Cerro ([Fa85], [Fa86]) concernant le principe de résolution en logique modale et la mise au point du système MOLOG, qui est une extension de PROLOG permettant les opérateurs modaux.

3.2 Logiques et formalismes numériques. Les méthodes considérées jusqu'ici reposent sur une approche purement symbolique où les logiques introduites permettent la manipulation de symboles formels, sans qu'aucune connotation numérique n'entre en jeu. Nous voulons maintenant, pour terminer, indiquer brièvement quelques approches utilisant des formalismes numériques dans le traitement d'informations incertaines et la modélisation du raisonnement (l'exposé [BK86] constitue une excellente introduction au sujet). L'idée est d'attribuer un certain poids numérique (habituellement un nombre réel dans l'intervalle $[0, 1]$) à l'information dont on dispose et de mettre en place un "calcul" permettant d'en tirer un poids pour les conclusions. Ainsi, si un agent croit un énoncé α , il arrivera souvent qu'il ne le croira pas de façon absolue mais jusqu'à un certain *degré*. En quantifiant ce degré de croyance, on pourra ainsi rendre plus explicite le statut de α . Il faut remarquer qu'une telle approche ne relève pas des mêmes préoccupations que les logiques non monotones: dans ce dernier cas, en effet, quand une chose est crue, elle est crue de façon absolue, sans nuances, même si elle est susceptible d'être éventuellement rétractée.

Une façon naturelle d'attribuer un poids à une information donnée est d'utiliser la notion de probabilité. Nilsson (voir [GN87]) montre que l'on peut préciser le concept de *probabilité d'un énoncé* en faisant intervenir l'idée de mondes possibles. Un énoncé α pouvant être soit vrai soit faux, on peut donc imaginer deux classes de mondes possibles, une première, M_1 , contenant les mondes où α est vrai et l'autre, M_2 , ceux où il est faux. Le monde actuel doit être dans l'une de ces classes mais nous pouvons ignorer laquelle. Une façon de modéliser cette incertitude est de donner une probabilité p_1 (selon une certaine mesure de probabilité) pour que le monde actuel soit dans M_1 et une probabilité $p_2 = 1 - p_1$ pour qu'il soit dans M_2 . On pourra alors dire que p_1 et p_2 représentent la probabilité que α soit respectivement vrai et faux. Une telle vision des choses recoupe certains aspects des logiques multivalentes dont il a été question plus haut. Dans le cas de plusieurs énoncés, on obtiendra plusieurs classes M_i de mondes possibles et pour chacune une certaine probabilité p_i que le monde actuel s'y trouve. Nilsson construit une *logique probabiliste* reposant sur cette façon d'attribuer des probabilités à des énoncés et introduit, de manière analogue à la notion de conséquence logique, une notion de conséquence probabiliste.

Le calcul des probabilités classique, qui à certains égards fournit un cadre trop rigide pour rendre compte de tous les aspects de l'incertitude (voir à ce propos les arguments présentés dans [DP85]), n'est pas le seul formalisme qui ait été utilisé en vue de modéliser l'imprécision dans les bases de connaissances. Ainsi, suite aux travaux de Dempster cherchant à traduire une telle imprécision via les concepts de probabilités supérieures et inférieures, Shafer [Sh76] a introduit, dans le cadre d'une *théorie de l'évidence*, la no-

tion de *fonctions de croyance* permettant de formuler des degrés de plausibilité et de crédibilité de certaines informations. Cette approche a suscité un intérêt marqué en IA (voir [KL86] et [Sm88] pour des résultats récents dans cette direction). Une autre avenue de recherche prend son origine dans la *théorie des possibilités* formulée par Zadeh [Za78]; Dubois et Prade ont apporté une contribution majeure à cette approche (voir [DP85], [DP88]) en formalisant une notion de *logique possibiliste*. Ils introduisent des mesures de confiance qui leur permettent d'associer à tout énoncé α un degré de possibilité $\Pi(\alpha)$ et un degré de nécessité $N(\alpha)$: ces degrés, compris dans l'intervalle $[0, 1]$, peuvent être vus comme des transcriptions numériques d'opérateurs correspondants en logique modale. $N(\alpha) = 1$ traduit le fait que α est nécessairement vrai tandis que $\Pi(\alpha) = 0$ signifie qu'il n'est pas possible que α soit vrai. Ces mesures satisfont certaines propriétés de base telles $\Pi(\alpha) \geq N(\alpha)$ (un énoncé est possiblement vrai avant d'être nécessairement vrai) ou encore $\Pi(\alpha) = 1 - N(\neg\alpha)$ (α impossible équivaut à $\neg\alpha$ nécessairement vrai). Dubois et Prade montrent comment construire des calculs à partir de ces mesures, permettant ainsi soit d'obtenir une logique monotone analogue à la logique probabiliste de Nilsson, soit de reproduire un raisonnement non monotone.

Dans [DP88], Dubois et Prade distinguent entre une logique de l'*incertitude*, effectuant un raisonnement précis sur de l'information imprécise, et une logique du *vague*, où le mode de raisonnement lui-même est non précis. Contrairement aux trois formalismes numériques précédents (logique probabiliste, logique de l'évidence, logique possibiliste), la *logique floue* s'intéresse à des raisonnements "flous" sur de l'information précise. Tout comme la théorie des possibilités, la logique floue repose sur la notion d'*ensemble flou* duc à Zadeh, c'est-à-dire d'ensemble dont la fonction caractéristique prend ses valeurs dans tout l'intervalle $[0, 1]$ (et non pas seulement aux extrémités $\{0, 1\}$ comme dans le cas ordinaire). On y retrouve donc des prédicats flous, c'est-à-dire mal définis, (par exemple "être grand") et même des quantificateurs flous ("la plupart", "beaucoup", etc.) à partir desquels on peut effectuer des raisonnements formels (voir [DP88] ou [ID90] pour plus de détails). Certains travaux récents illustrent comment cette approche est susceptible d'être mise en application dans un contexte à la PROLOG (voir le FPROLOG de [MB87] ou encore le FLOULOG de [Ga87] dans lequel se trouve incorporée une notion de non pertinence).

Certaines des approches précédentes relevant d'attitudes fort différentes, il n'est donc pas étonnant de constater que l'évaluation de leurs mérites est souvent malaisée et donne lieu à des points de vue divergents. Par delà la controverse classique en théorie des probabilités divisant les "fréquentistes", qui considèrent les probabilités comme les limites de fréquences relatives observées, et les "subjectivistes" ou "bayesiens", qui les voient comme des mesures d'un sentiment d'incertitude, il y a surtout débat concernant le caractère essentiel ou non de tel ou tel formalisme. La littérature récente contient de nombreux exemples de telles prises de position, comme en témoignent par exemple plusieurs textes du recueil [KL86].

4. Conclusion. On rapporte que Leibniz était plutôt optimiste quant au délai requis pour mener à terme son projet de mécanisation du raisonnement:

"I think that a few selected men could finish the matter in five years. It would take them only two however to work out by an infallible calculus the doctrines most useful for life, those of morality and metaphysics." (Cité dans [Sh83], p. 6)

Et pourtant non seulement est-il impossible, comme il a déjà été souligné plus haut, de réaliser le projet de Leibniz en tant que tel, mais même dans des réalisations partielles semble-t-il extrêmement difficile d'arriver à des résultats concrets probants. Les problèmes entourant l'acquisition et la représentation de la connaissance ainsi que l'inférence automatisée ont provoqué le développement de théories nombreuses, variées et de plus en

plus complexes. Le survol qui précède visait, dans les limites du cadre restreint de ce travail, à donner au lecteur un aperçu de la richesse du sujet.

S'il est un domaine où il ne saurait être question de soutenir une position de monisme quant aux fondements théoriques, c'est bien celui dont nous avons traité ici. Les nombreux formalismes que nous venons d'esquisser ont été conçus chacun en vue de traduire certains aspects du raisonnement de sens commun. Autant il est important de favoriser l'émergence de nouvelles avenues visant à jeter un éclairage meilleur sur certaines facettes du problème, autant il est également nécessaire de disposer d'études de synthèse favorisant la compréhension globale et la comparaison des différents formalismes en cause. De nombreux ouvrages de cette nature sont fort heureusement disponibles grâce auxquels le lecteur peut approfondir sa connaissance du sujet et parmi les plus récents, on peut retenir [Ga86], [GN87], [ID90], [MW8-], [Pa90], [Ra88], [SM88], [So89] et [Th8-], ouvrages que nous aimerions commenter brièvement.

On trouvera dans [Ga86] et [MW8-] (principalement dans le Volume 2 portant sur la structure d'inférence) des présentations de la logique classique qui prennent en considération un passage éventuel à l'automatisation de la déduction; dans les deux cas, c'est la technique du calcul des séquents qui est à la base des formalismes mis en place. Le livre [GN87], tout comme [Ra88], se veut un texte de base sur les divers rôles de la logique en IA; prenant la logique classique comme point de départ, ces deux ouvrages introduisent de façon systématique différentes logiques non classiques. ([GN87] est utile en particulier par son abondante bibliographie.) Le recueil de textes [ID90] contient plusieurs exposés de fond traitant de toutes les logiques non classiques mentionnées plus haut; il en est de même du recueil [SM88], où chaque article est de plus accompagné de quelques critiques sur son contenu ainsi que de répliques des auteurs (le contenu de [SM88] ayant fait l'objet de discussions lors d'ateliers réunissant tous les auteurs). L'ouvrage [Th8-] veut faire une présentation systématique des concepts et méthodes de l'IA en prenant comme ligne directrice la logique sous toutes ses formes; un troisième volume est en préparation, consacré à des exemples de mise en oeuvre de la logique dans des applications de l'IA. Un tel point de vue pratique se retrouve dans [Pa90]: après avoir passé en revue les principaux systèmes de représentation de la connaissance que nous avons décrits ici, ce livre s'intéresse aux aspects des systèmes experts reliés tant à la communication et à la perception qu'à l'acquisition ou à l'organisation de connaissances. Enfin l'étude comparative [So89], du collectif Léa Sombé, est particulièrement précieuse par l'analyse qu'elle présente de plusieurs systèmes d'inférence non classique pour l'IA; les auteurs y ont identifié une série de critères à l'aide desquels ils comparent les différents formalismes sous étude (voir dans [So89] les pp. 18-19 de même que le tableau synthétique de la page 199).

N.D.L.R. A l'occasion, la revue publie des articles de survol. Etant donnée l'effervescence de ce sujet ces dernières années, cet article nous apparaît utile et pertinent.

BIBLIOGRAPHIE

- [Ba77] Barwise J., "Handbook of Mathematical Logic," North-Holland, 1977.
- [BF85] Barwise J. et Feferman S., "Model-Theoretic Logics," Springer-Verlag, 1985.
- [Be89] Besnard P., "An Introduction to Default Logic," Springer-Verlag, 1989.
- [BK86] Bhatnagar R.K. et Kanal L.N., *Handling uncertain information: a review of numeric and non-numeric methods*, in "[KL86]," pp. 3-26.
- [Bi86] Bibel W., *Automated inferencing*, J. Symbolic Comput. 1 (1985), 245-260.
- [BL85] Brachman R.J. et Levesque H.J., "Readings in Knowledge Representation," Morgan Kaufman, 1985.

- [Cl78] Clark K., *Negation as failure*, in "Logic and Data Bases," (Aussi in [Gi87], pp. 311-325), Plenum, 1978, pp. 293-322.
- [Da83] Davis M., *The prehistory and early history of automated deduction*, in "Automation of Reasoning vol 1: Classical Papers on Computational Logic, 1957-1966," Springer-Verlag, 1983, pp. 1-28.
- [dS88] de Swart H.C.M. et Ophelders W.M.J., *Another tableau based automated theorem prover for classical logic*, Bull. EATCS 35 (1988), 185-197.
- [DP85] Dubois D. et Prade H., "Théorie des possibilités: Applications à la représentation des connaissances en informatique," (2e édition, 1987) (Traduction anglaise: "Possibility Theory: An Approach to Computerized Processing of Uncertainty," Plenum, 1988), Masson, 1985.
- [DP88] ———, *An introduction to possibilistic and fuzzy logics*, in "[SM88]," pp. 287-315.
- [Et88] Etherington D.W., "Reasoning with Incomplete Information," Pitman, 1988.
- [Fa85] Fariñas del Cerro L., *Resolution modal logic*, Logique et Analyse 28 (1985), 153-172.
- [Fa86] ———, *MOLOG: A system that extends PROLOG with modal logic*, New Generation Comput. 4 (1986), 35-50.
- [GG8-] Gabbay D. et Guenther F., "Handbook of Philosophical Logic, Vol. 1: Elements of Classical Logic;" "Vol. 2: Extensions of Classical Logic;" "Vol. 3: Alternatives to Classical Logic," Reidel, 1983 (Vol. 1), 1984 (Vol. 2), 1986 (Vol. 3).
- [Ga87] Gagné J.-R., "Langage de programmation avec logique floue et sous-ensembles flous," Thèse de maîtrise, Université Laval, 1987.
- [Ga86] Gallier J.H., "Logic for Computer Science: Foundations of Automatic Theorem Proving," Harper & Row, 1986.
- [GN87] Genesereth M.R. et Nilsson N.J., "Logical Foundations of Artificial Intelligence," Morgan Kaufmann, 1987.
- [Gi87] Ginsberg M.L., "Readings in Nonmonotonic Reasoning," Morgan Kaufman, 1987.
- [Ha82] Hatcher W.S., "The Logical Foundations of Mathematics," Pergamon, 1982.
- [HC68] Hughes G.E. et Cresswell M.J., "An Introduction to Modal Logic," Methuen, 1968.
- [KL86] Kanal L.N. et Lemmer J.F., "Uncertainty in Artificial Intelligence," North-Holland, 1986.
- [Kr63] Kripke S., *Semantical analysis of modal logic*, Z. Math. Logik Grundlag. Math. 9 (1963), 67-96.
- [ID90] Iturrioz L. et Duchaussoy A., "Modèles logiques et systèmes d'intelligence artificielle," Hermès, 1990.
- [Le77] Leibniz G.W., "Préface à la Science générale (1677)," in "Opuscules et fragments inédits de Leibniz, extraits des manuscrits de la Bibliothèque royale de Hanovre", Félix Alcan, 1903, pp. 153-157.
- [MW8-] Manna Z. et Waldinger R., "The Logical Basis for Computer Programming Vol. 1: Deductive Reasoning;" "Vol. 2: Deductive Systems," Addison-Wesley, 1985 (Vol. 1), 1990 (Vol.2).
- [MB87] Martin T.P., Baldwin J.F. et Pilsworth B.W., *The implementation of FPROLOG - A fuzzy PROLOG interpreter*, Fuzzy Sets and Syst. 23 (1987), 119-129.
- [Mc58] McCarthy J., *Programs with common sense*, in "[BL85]," 1958, pp. 299-307.
- [Mc80] ———, *Circumscription - A form of non-monotonic reasoning*, (Aussi in [Gi87], pp. 145-152), Artificial Intelligence 13 (1980), 27-39, 171-172.
- [Mc82] McDermott D., *Non-monotonic logic II: non-monotonic modal theories*, J. Assoc. Comput. Mach. 29 (1982), 33-57.
- [MD80] McDermott D. et Doyle J., *Non-monotonic logic I*, (Aussi in [Gi87], pp. 111-126), Artificial Intelligence 13 (1980), 41-72.
- [Me64] Mendelson E., "Introduction to Mathematical Logic," (3e éd., Wadsworth, 1987), Van Nostrand, 1964.
- [Me89] Meseguer J., *General logics*, in "Logic Colloquium '87," Studies in Logic #129, North-Holland, 1989, pp. 275-329.
- [Mo82] Moore R.C., *The role of logic in knowledge representation and commonsense reasoning*, (Aussi in [BL85], pp. 335-341), in "Proc. AAAI-82," Morgan Kaufmann, 1982, pp. 428-433.
- [Mo88] ———, *Autoepistemic logic*, in "[SM88]," pp. 105-127.
- [Ne86] Nerode A., *Applied Logic*, in "The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics," Springer-Verlag, 1986, pp. 127-163.
- [Pa90] Patterson D.W., "Introduction to Artificial Intelligence and Expert Systems," Prentice-Hall, 1990.
- [Ra88] Ramsay A., "Formal Methods in Artificial Intelligence," Cambridge University Press, 1988.
- [Re78] Reiter R., *On closed world data bases*, (Aussi in [Gi87], pp. 300-310), in "Logic and Data Bases," Plenum, 1978, pp. 55-76.

- [Re80] ———, *A logic for default reasoning*, (Aussi in [Gi87], pp. 68–93), *Artificial Intelligence* 13 (1980), 81–132.
- [Re87] ———, *Nonmonotonic reasoning*, *Annual Rev. Comput. Sci.* 2 (1987), 147–186.
- [Re68] Rescher N., "Topics in Philosophical Logic," Reidel, 1968.
- [Ri89] Richards T., "Clausal Form Logic: An Introduction to the Logic of Computer Reasoning," Addison-Wesley, 1989.
- [Ro65] Robinson J.A., *A machine-oriented logic based on the resolution principle*, *J. Assoc. Comput. Mach.* 12 (1965), 23–41.
- [Sh76] Shafer G., "A Mathematical Theory of Evidence," Princeton University Press, 1976.
- [Sh83] Shepherdson J.C., *The calculus of reasoning*, in "Intelligent Systems: The Unprecedented Opportunity," Ellis Horwood, 1983, pp. 3–17.
- [Sh87] Shoham Y., *A semantical approach to nonmonotonic logics*, (Aussi in [Gi87], pp. 227–250, – version augmentée), in "Proc. 2nd Symposium on Logic in Computer Science," 1987, pp. 275–279.
- [Sh88] ———, "Reasoning About Change," MIT Press, 1988.
- [Sm88] Smets P., *Belief functions*, in "[SM88]," pp. 253–277.
- [SM88] Smets P., Mamdani A., Dubois D. et Prade H., "Non-Standard Logics for Automated Reasoning," Academic Press, 1988.
- [So89] Sombé L. (collectif), "Raisonnements sur des informations incomplètes en intelligence artificielle," (Version remaniée de: *Inférences non-classiques en intelligence artificielle: ébauche de comparaisons sur un exemple*, in "Intelligence artificielle," (Actes des journées nationales du PRC-GRECO, Toulouse, (1988) Teknea, (1988), pp. 137–230), Teknea, 1989.
- [Su83] Sundholm G., *Systems of deduction*, in "[GG83]," pp. 133–188.
- [Th8-] Thayse A. et al., "Approche logique de l'intelligence artificielle, Vol.1: De la logique classique à la programmation logique;" "Vol. 2: De la logique modale à la logique des bases de données," Dunod, 1988 (Vol.1), 1989 (Vol. 2).
- [Tu84] Turner R., "Logics for Artificial Intelligence," (Traduction française: "Logiques pour l'intelligence artificielle," Masson, 1986), Ellis Horwood, 1984.
- [vD83] van Dalen D., "Logic and Structure (2e éd.)," Springer, 1983.
- [vH67] van Heijenoort J., "From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931," Harvard Univ. Press, 1967.
- [Wa90] Wallen L.A., "Automated Deduction in Nonclassical Logics," MIT Press, 1990.
- [WN81] Webber B.L. et Nilsson N.J., "Readings in Artificial Intelligence," Morgan Kaufman, 1981.
- [WH83] Wos L. et Henschen L., *Automated theorem proving 1965-1970*, in "Automation of Reasoning. Vol. 2: Classical Papers on Computational Logic, 1967-70," Springer-Verlag, 1983, pp. 1–24.
- [WV87] Wos L. et Veroff R., *Resolution, binary: its nature, history and impact on the use of computers*, in "Encyclopedia of Artificial Intelligence," Wiley, 1987, pp. 892–902.
- [Za78] Zadeh L.A., *Fuzzy sets as a basis for a theory of possibility*, *Fuzzy Sets and Systems* 1 (1978), 3–28.
- [Za84] ———, *Coping with the imprecision of the real world*, *Comm. ACM* 27 (1984), 304–311.

B.R. Hodgson
 Département de mathématiques et de statistique
 Université Laval
 Québec, Canada
 G1K 7P4